CONNEXIONS™

The Interoperability Report

June 1988 Volume 2, No. 6

ConneXions - The Interoperability Report tracks current and emerging standards and technologies within the computer and communications industry.

In this issue:

ConneXions is published by Advanced Computing Environments, 480 San Antonio Road, Suite 100, Mountain View, CA 94040, USA. Phone: 415-941-3399.

ISSN 0894-5926

From the Editor

As mentioned in a previous issue of *ConneXions*, TCP/IP is very popular in Europe. This month we bring you a report from the Hannover CeBIT Fair which took place in mid-March.

One of the most complex issues in computer networking is routing. Ross Callon, Varda Haimo, and Marianne (Gardner) Lepp, all of BBN Communications Corporation give an overview of this topic.

There has been a flurry of activity relating to the development of standards for network management in the Internet community. The Internet Activites Board held a meeting in late March and subsequently issued RFC 1052 with their recommendations on how these activivies should be coordinated. We bring you an extract from this RFC, and will follow this up with reports from the various network management groups in future issues.

The topic of transition from TCP/IP to ISO/OSI has already been adressed in several articles in *ConneXions*. This month we look at the DoD's transition strategy. The article is by Gladys Reichlen and Phill Gross of the Mitre Corporation.

In our series “Improving Your TCP”, Craig Partridge looks at the basic tools of the trade in an article entitled “Look under the hood!”.

You may have seen references to IDEAS in recent Internet discussions. The Internet Design, Engineering and Analysis Series (IDEAS) are draft documents of the Internet Engineering Task Force (IETF). IDEAS are generally contributed by IETF Working Groups on short- and mid-term issues in network, internetwork and protocol engineering. However, thoughtful papers from any responsible source on related issues will be considered. The IETF chairman is the nominal editor of the series and can be reached by sending electronic mail to gross@gateway.mitre.org.

Since this is a draft document series, with the goal of giving wider exposure to the technical issues, minimal editing has been done for form or style. IDEAS are *not* standards and should *not* be utilized in RFPs or product announcements. The basic format is the same as that of the RFCs, since some IDEAS may ultimately be submitted as RFC's. Constructive comments on all IDEAS are encouraged. Such comments can be addressed directly to the author or, preferably, can be sent to: ietf@venera.isi.edu. The IDEAS are available from the IETF: directory on host SRI-NIC.ARPA.

TCP/IP at the Hannnover CeBIT Fair

by Horst Clausen, University of Kansas

Hannover, Federal Republic of Germany, March 16-23, 1988 -- One of the world's biggest industrial fairs and exhibits is the annual Hannover Messe which takes place every spring. Not only is it an exhibit of Germany's industrial products but also a showplace for all industrial nations and products. The data processing and telecommunications industries have grown over the years to a size which forced the organizers to split this part off and run a separate fair about one month prior to the other industrial products exhibition. This has happened now for three years and this fair has been called the "CeBIT Messe". The CeBIT Fair is an annual event and it is competing for first place against the "Systems" which takes place in Munich every second year in the fall; so far it has not been decided which event will be the most important computer and communications exhibit in Germany.

ISO/OSI

It is well known that there is a lot of interest in the ISO OSI standards and protocols in Europe, not only among government agencies and international institutions but also among large users. Since telecommunications is strictly regulated and controlled by the PTTs, any form of wide-area network has to adhere to the standards and protocols dictated by the PTTs. In most countries this boils down to X.25 and/or X.21 and to leased lines. Since the tariffs for leased lines are being raised continuously in order to make the Public Data networks more competitive, large users of telecommunications in Europe have to look for solutions which are different from the ones found in the United States.

ISDN services

With the upcoming introduction of ISDN services the picture will change again. In this context it is interesting to note that the German Bundespost has announced a tariff for ISDN services which at the moment is applicable only to the two pilot systems in the cities of Mannheim and Stuttgart; they have also publicly announced that ISDN services will be installed in all major cities and between all major cities in West Germany within the next 3 years. At this time it is, however, totally unclear if and how the current packet switching X.25 network, Datex-P, will be made accessible to ISDN users through the signalling channel; at present the only way is to dial a B-channel connection.

Local Area Networks

In the realm of local area networks, Ethernet products lead the market; the majority of the systems are US products which are either offered by local distributors or by international branches of US companies. People who install LANs usually have a set of problems to solve and the LAN is part of their solution; the other part is getting the various pieces to talk to each other. A typical example of this is the currently very fashionable CAD/CIM area where people are integrating various systems into one distributed system. And in this area there is only one protocol family that is being used: TCP/IP. Although there is lots of talk about OSI protocols, the fact remains that most of todays hardware and software products talk TCP/IP and that you can get it for almost any system for a reasonable price. People also know that these protocols do work and what efficiency, throughput and delay to expect.

Multi-vendor demonstration

In this respect the real break-through was a demonstration at the previous Systems '87 in Munich last fall when some 30 vendors showed internetting with TCP/IP across one common Ethernet cable. This demonstration has made a tremendous impact and everyone who needs a product now is looking at TCP/IP.

Applications

This does *not* mean that TCP/IP and the applications such as Telnet, SMTP and FTP all of a sudden have become beloved protocols among the people who are pushing towards "Open Systems" and standards. Anybody who believes that the OSI protocols will not be successful in the long run is fooling himself. However, contrary to what the network wizards believe, OSI will not be built from the bottom up by implementing the proper subnetworks with ISO-IP and TPn layers first, but from the top and for specific applications. Examples are the X.400 electronic message systems and other applications in the industrial area. This is where the users have a true interest and also an investment, and they are not particularly interested in which way the bits get to the "other side"; basically this includes all lower level protocols and everything that is "hidden" inside the systems software.

Transitioning to OSI

Users will probably be very happy when they learn about the ISODE approach which will allow them to still go out and buy TCP/IP interfaces and software and develop or run a complete OSI protocol suite on top of it. This is done by emulating a TP0 interface on top of TCP and hiding the TCP/IP specifics inside that interface. This is a tremendously practicable approach and it beautifully integrates the strong parts of each protocol family. And last not least - it works!

Coming back to CeBIT: TCP/IP has made its inroads into Europe and it will continue to do so over the next years. The majority of the products are of US origin, there are few companies in Europe building their own TCP/IP hardware or software. The availability of hosts with PTT approved X.25 interfaces (e.g. SUNs) makes interconnecting LANs via X.25 with TCP/IP possible - at rather low cost and effort. ISDN has now entered the scene and will make interesting alternatives to interworking LANs possible; vendors better be prepared to offer true 64 kbit/sec interfaces for ISDN soon. The push towards OSI standards comes mostly from the applications side and not from the networks. And any future solution will have to guarantee a smooth transition from the current TCP/IP protocols to whatever is going to replace them. When is this going to happen? - when all FORTRAN code has been replaced by Ada code.

HORST D. CLAUSEN is a Professor of Computer Engineering at the University of Kansas. Previously he spent 5 years at DFVLR - the German equivalent of NASA - working on the SATNET experiment. Before joining DFVLR he participated in the development of the Ada programming language. He holds a PhD in Theoretical Physics from the University of Graz.

Routing in an Internetwork Environment

by Ross Callon, Varda Haimo, and Marianne (Gardner) Lepp, BBN Communications Corp.

Introduction and overview

The Internet provides a set of routing problems different from those of other network environments. The Internet is a *very* diverse environment under the control of multiple administrations, which utilizes equipment built by a wide variety of vendors. The design of an Internet routing algorithm must accommodate this fact. Because multiple administrations control the Internet, changes are hard to make and incorrect tuning can occur. In addition, the routing algorithm will be used in an environment in which the DoD and ISO internet protocols are used in parallel, with some gateways providing a dual IP function.

Autonomous Systems

Internet gateways are partitioned into *Autonomous Systems* [1]. This allows different parts of the Internet to be administratively separate, and allows gateways in different Autonomous Systems to use different routing protocols internally. For example, the BBN LSI-11 gateways use the GGP protocol, the newer BBN Butterfly gateways use a version of SPF, and many of the gateways in the NSFNET use RIP [2]. Gateways in different Autonomous Systems exchange reachability information via the *Exterior Gateway Protocol* (EGP) [1,3]. A similar approach is used in the OSI environment, where gateways (or "Intermediate Systems") as well as networks and hosts are split into "routing domains" [4-6].

This article considers issues in the design of a routing protocol for use within a single autonomous system or routing domain. Routing between domains is greatly complicated by the numerous administrative restrictions that may be imposed, and is beyond the scope of this article. The requirements for routing between multiple Autonomous Systems or routing domains in discussed in [7].

There are a number of issues that must be addressed in discussing the question of which techniques are appropriate for intra-Autonomous System routing in the Internet. Among the most important issues are the choice of algorithm, what levels of hierarchy the routing algorithm should have, how dynamic routing should be, the choice of metrics, and how routing information should be distributed. Each of these is discussed below.

Routing algorithms

There are two classes of commonly employed algorithms which choose shortest paths (with respect to a particular metric). We call *distance-vector* algorithms the class of algorithms in which a node sends to its neighbors a vector of distances (its routing table) and *link-state* algorithms those in which a node floods to all nodes the metric associated with its adjacent links. With distance-vector methods a node uses its neighbors' routing tables to determine routes. This apparently gives distance-vector algorithms an overhead advantage as compared to link-state algorithms. However, the vector of information shared between neighbors can be quite large. If a gateway has many neighbors this large update can be sent almost as many times as the small update from link state algorithms. This problem can be addressed by restricting the number of neighbors who receive updates to a number smaller than all a gateway's internet neighbors.

In addition, since each vector-state update often results in a chain of updates, this can constitute a de facto form of flooding. In general, however, distance-vector algorithms scale well.

Another apparent advantage for distance-vector routing is that only a small amount of CPU is used -- one addition and one comparison for each neighbor. In contrast, the CPU used by link-state algorithms scales with the square of the number of nodes. This advantage is illusory, however, because the potentially slow convergence of distance-vector algorithms may force nodes to repeat these simple calculations many times before routes stabilize. Link-state algorithms use more CPU per computation, but they calculate new routes directly rather than converging towards the solution, thereby offsetting their apparent additional CPU requirements. In addition, link-state algorithms can reduce the amount of needed computation by doing incremental updates to the routing tables [8].

Routing loops

The most serious disadvantages to distance-vector based algorithms are that they can form *routing loops*, and they can be slow to converge, leading to route instability and increased overhead. There are a large number of fixes that have been implemented and/or proposed to reduce this problem [10-14]. However, these numerous fixes to distance-vector algorithms only reduce the seriousness of these problems -- they do not eliminate them. In addition, since no global information is maintained, and since every node maintains a different set of information which is calculated from the information received from neighbors, errors in routing algorithm performance tend to be hard to identify. These convergence problems are discussed in greater detail in [15].

In contrast, link-state methods require that information about the state of the whole network be distributed across the net, resulting in a control traffic burden in large networks. The benefit of the link-state method is that the gateways maintain consistent views of the network, thus precluding problems such as looping and slow adjustment to changes in network conditions. In addition, since each node maintains an identical database, misbehaving nodes are somewhat easier to detect. Finally, link-state algorithms are easier to use with Type-of-Service Routing [17]. The relative advantages of each of these approaches is discussed in greater detail in [15, 16].

Hierarchical routing

The level of hierarchy of the Internet routing algorithm is an issue of exceptional importance given the rate of growth of the Internet. Both link-state and distance-vector algorithms will have problems in very large systems -- link-state because of the amount of control traffic it engenders, and distance-vector because of its convergence problems. If Autonomous Systems grow very large, some level of hierarchy must be included in Internet routing to reduce the amount of information that must be processed at each node.

The main issues for hierarchical routing are determining where and how to define the levels of hierarchy, how to adjust to changes in topology (particularly changes that partition an area), how to compute routes, and how to propagate the routing information (including the information used to calculate routes, and/or the routes themselves).

continued on next page

Routing in an Internetwork Environment *(continued)*

There are a large number of ways to arrange networks into hierarchies. For example, "quasi-hierarchical" methods such as Kamoun-Kleinrock clustering [18] and Landmark Routing [19] use a hierarchy to summarize information, but combine detailed local information and summarized remote information into a single level of routing calculations. "Pure Hierarchical" methods separate a network or internet into multiple levels, and perform separate routing calculations at each level. Higher level routing calculations may compress lower level information in one of several ways. For example, higher level routing calculations may consider each underlying area to be a single node [20]. Alternatively, each area may be condensed into several nodes, such as one for each boundary with a neighboring area [21]. The manner in which information is hierarchically summarized has an effect on both the amount of routing information which is necessary to perform routing calculations, and on the quality and stability of the routes which are chosen.

Hierarchical routing algorithms can be used in systems which are very large, but there are costs. Since routing information is summarized, routes will not be as good as those found by a flat routing scheme. However, our work on SURAN included simulations which showed that the difference in routes is generally not significant [20]. Because of the problems with area determination and area partitions, hierarchical algorithms are more complicated than flat routing algorithms. There are also tradeoffs between the complexity of the algorithm and performance. For example, Kamoun-Kleinrock clustering or Landmark Routing will minimize the complexity of a hierarchical scheme, but requires the use of a distance vector routing algorithm which suffers from other problems.

Dynamic versus static routing algorithms

Routing algorithms can vary in responsiveness from using static paths to having paths change dynamically in response to the congestion state of resources. A static algorithm finds its routes and its back-up routes from tables provided to the gateway at configuration time. Such tables cannot be completely responsive in the face of failures since back-up routes may overlap on the failed resource. Furthermore, as the physical topology changes, every gateway must have its tables updated. Another serious problem with static routing is the difficulty of configuring routing tables correctly. As the internet grows the configuration problem rapidly worsens, thus introducing an increasing risk of mis-configuration. Static routing tables provide very stable routing, however.

A dynamic routing algorithm adapts to changing conditions. It can respond automatically to topology changes, or, more dynamically, it can also respond to congestion. However, dynamics can be a double-edged sword. Care must be taken to avoid routing oscillations and their attendant costs in control overhead, CPU, and congested resources. In a large and heterogeneous network environment such as the Internet, one must be very careful not to design a routing algorithm that responds to network conditions at a rate faster than the rate at which relevant information can reach decision points in the network.

For instance, a routing algorithm that responds to congestion after the congestion has already dissipated is not well-designed. These considerations are particularly important in the DoD Internet where control information must travel long distances, and where there is a large range of network types.

Another major advantage of dynamic routing schemes is that they allow some degree of automatic configuration of a network or internet. For example, dynamic routing may automatically incorporate new connections into its routes and automatically eliminate failed connections from its routes. The hierarchical scheme designed for the SURAN networks automatically configures packet radios into a layered hierarchy. Similarly, Landmark Routing allows for a high degree of self-configuration.

Choice of metrics

Metrics provide the criteria by which routing algorithms choose between alternate routes. As such, the appropriate choice of metrics for an internet routing algorithm is determined in large part by the requirements for the algorithm. Some possible choices are *hop count*, *administrative distance*, *capacity*, or *delay*. A metric that is fixed will yield stable routes, and will also permit routing to respond to outages. A metric that measures congestion will yield routes which change to avoid congestion, but this may cause routing oscillations if the metric is poorly chosen. The metric should be chosen in the context of the requirements for routing, and with considerations of consistent and effective algorithm design at the fore. The choice of routing metrics used to determine routes will have a major impact on both the quality of routes and the stability of the algorithm.

Distribution of routing information

Routing can be computed with a distributed computation or it can be computed at a centralized routing node (an "oracle") and distributed on request to gateways. With a distributed algorithm, routes can be computed jointly by each gateway so that only the destination is needed for forwarding gateways to identify the next hop. Alternatively, the source gateway can compute the route and source-route data.

With centralized routing the oracle makes all routing decisions and then announces them to the rest of the network. Depending upon the adaptivity of the routing algorithm, this central router needs information from the network on which to base its routing decisions. Thus it is necessary to have very good communication and sufficient capacity between the oracle and the rest of the network. The advantages are that routes chosen by a central router will be consistent, and that other network resources need not be expended on processing routing choices. The disadvantages are that the network is very vulnerable to congestion and communications problems at the central routing site or sites.

With distributed routing there are two main choices for determining paths. The source can determine routes, and then include them with the packet or install routes at forwarding gateways. The alternative is to allow each gateway to determine what the appropriate next hop should be based on the destination of the packet (and perhaps on other factors such as packet type). The latter method is used by most single-path routing algorithms.

continued on next page

Routing in an Internetwork Environment *(continued)*

With multipath routing, consistency of routes is hard to maintain, making source routing attractive. Source routing would increase the size of the packet and use more memory and CPU in the gateways. Alternative encodings of the source route are possible, but this would require redesign of the Internet Protocols or definition of an IGP gateway-to-gateway header. The advantages of a source-routing scheme, such as no-looping, no repeated computation, and ease of encoding multiple paths in a multipath routing scheme, may be insufficient to justify the changes that it would entail.

References

[1] Linda J. Seamonson and Eric Rosen, "Stub Exterior Gateway Protocol," RFC 888, Jan 1984.

[2] C. Hedrick, "Routing Information Protocol," Rutgers University, IDEA 004, Nov 1987

[3] Dave Mills, "Exterior Gateway Protocol Formal Specification," RFC 904, Apr 1984.

[4] "OSI Routing Framework," ISO TC97/SC6/N4616, June 1987.

[5] Tassos Nakassis, "Basic Issues of Routing Between Routing Domains," NBS, Sept 1987.

[6] "Intermediate System to Intermediate System Intra-Domain Routing Exchange Protocol," source USA, ISOTC97/SC6/N4945, Oct 1987.

[7] Ross Callon (editor) "Requirements for Inter-Autonomous Systems Routing," IDEA 007, Jan 1987.

[8] J. McQuillan, I. Richer, and E. Rosen, "ARPANET Routing Algorithm Improvements, First Semiannual Technical Report," BBN Report 3803, Apr 1978.

[9] E. Rosen, J. Mayersohn, P. Sevcik, G. Williams, and R. Attar, "ARPANET Routing Algorithm Improvements, Volume 1," BBN Report 4473, Aug 1980.

[10] J. Jaffe and M. Moss, "A Responsive Distributed Routing Algorithm for Computer Networks," IEEE Transactions on Communications, vol. COM-30, pp. 1758-1762, Jul 1982.

[11] P. M. Merlin and A. Segall, "A failsafe distributed routing protocol," EE PUB No. 313, Dept. of EE, Technion - Israel Institute of Technology, Haifa, Israel (1978).

[12] P. M. Merlin and A. Segall, "A failsafe distributed routing protocol," IEEE Trans. Comm. Com-27, No. 9 (1979) pp 1280-1287.

[13] S. Toueg, "A minimum-hop path failsafe and loop-free distributed algorithm," IBM Research Report RC 8530 (1980).

[14] J. Garcia-Luna-Aceves, "A New Minimum-Hop Routing Algorithm," SRI International, 1987.

[15] R. Callon, "A Comparison of 'Link State' and 'Distance Vector' Routing Algorithms," IDEA 002, Nov 1987.

[16] L. Bosack, "A Further Comparison of 'Link State' and 'Distance Vector' Routing Algorithms," cisco Systems, Feb 1988.

[17] M. Gardner, I. Loobeek, and S. Cohn, "Type-of-Service Routing: Preliminary Design," BBN Report 6195, June 1986.

[18] L. Kleinrock and F. Kamoun, "Hierarchical Routing for Large Networks," Computer Networks, Vol 1, No. 3, Jan 1977.

[19] Paul F. Tsuchiya, "Landmark Routing: Architecture, Algorithms, and Issues," MTR-87W00174, Sept 1987.

[20] R. Callon and G. Lauer, "Hierarchical Routing for Packet Radio Networks," BBN Report 5945, SRNTN No. 31, June 1985.

[21] A. Khanna, "Large Network Routing Study: Final Report," BBN Report 6267, Oct 1986.

This work was supported by the Defense Advanced Research Project Agency, under contract number F29601-87-C-0086. Distribution is unlimited

ROSS CALLON received his B.Sc. in Mathematics from the Massachusetts Institute of Technology in 1973, and his M.Sc. in Operations Research from Stanford University in 1977. He has been with BBN since 1980, and is currently the Internet Architect for the Gateway Development group. He is concerned with a variety of design issues relating to the internet architecture, including internet routing, design of high speed gateways, congestion control, gateway monitoring and control, and protocol standardization. Mr. Callon is a member of the Internet Engineering Task Force, the IETF Open Routing Working Group, and ANSI Task Group X3S3.3.

VARDA HAIMO received her Ph.D. in Applied Mathematics from Harvard University in 1984. She spent a year as a Postdoctoral Research Fellow at the Harvard Business School, after which she came to BBN Communications. At BBN Varda manages the Algorithm Development, Simulation and Modelling group in the Network Analysis Department. She has worked extensively in the area of routing in packet-switched networks: participating in the design, modelling and simulation of a new Type-Of-Service routing algorithm for the DDN.

MARIANNE LEPP received her Ph.D in mathematics from the University of Wisconsin, Madison in 1975, and held academic positions at North Carolina State University and Worcester Polytechnic Institute. For the past 4 years she has worked at BBN Communications on a variety of networking projects including network design and design tool and algorithm development, performance analysis, and protocol development for both BBN packet-switched networks and the DoD Internet. Among her key projects was the design of a Type-Of-Service routing algorithm with loadsharing for BBN packet-switched networks. In addition, she is a member of the Internet Engineering Task Force, chaired the EGP3 working group and is a member the Open Routing working group, both part of the IETF.

IAB Recommendations for the Development of Internet Network Management Standards

by Vint Cerf, NRI

Introduction

This article is an extract from RFC 1052 and is intended to convey to the Internet community and other interested parties the recommendations of the Internet Activities Board (IAB) for the development of network management protocols for use in the TCP/IP environment. RFC 1052 does *not*, in and of itself, define or propose an Official Internet Protocol. It does reflect, however, the policy of the IAB with respect to further network management development in the short and the long term.

Background

At the IAB meeting on 21 March, 1988 in videoconference, the report of the Ad Hoc Network Management Review Committee was reviewed. The recommendations of the committee were endorsed by the IAB and direction given to the chairman of the Internet Engineering Task Force to take the necessary steps to implement the recommendations.

The IAB expressed its gratitude for the efforts of the HEMS, SNMP and CMIP/CMIS working groups and urged that parties with technical interest in the outcome of the network management working groups convey their ideas and issues to the relevant working group chairmen.

Three groups

The IETF chairman was directed to form two new working groups, one of which would be responsible for the further specification and definition of elements to be included in the Management Information Base (MIB). The other would be responsible for defining extensions to the Simple Network Management Protocol (SNMP) to accommodate the short-term needs of the network vendor and operator communities. The longer-term needs of the Internet community are to be met using the ISO CMIS/CMIP framework as a basis. A working group of the IETF exists for this work and would continue its work, coordinating with the two new groups and reporting to the IETF chairman for guidance.

The output of the MIB working group is to be provided to both the SNMP working group and the CMIS/CMIP ["Netman"] working group so as to assure compatibility of monitored items for both network management frameworks.

Sumary of ad hoc meeting

On 29 February 88, an ad hoc committee was convened to review the network management options for the Internet in particular and the TCP/IP protocol suite in general. This meeting was called at the request of the Internet Activities Board in the course of exercising its responsibilities to the Federal Research Internet Coordinating Council (FRICC) and by the Mitre Corporation as a consequence of its work for the U.S. Air Force on the ULANA project.

At the conclusion of the one day meeting, it was agreed that the following recommendations be forwarded to the IAB chairman, Dr. David C. Clark, for consideration at the next IAB meeting scheduled for 21 March:

- In the short term, the Internet community should adopt and adapt the Simple Network Management Protocol (SNMP) for use as the basis of common network management throughout the system. (Rationale: The software is available and in operation.)

- In the longer term, the Internet research community and the vendors should develop, deploy and test a network management system based on the International Organization for Standardization (ISO) Common Management Information Services/Common Management Information Protocol (CMIS/CMIP). (Rationale: The Internet community can take the high ground in protocol development by virtue of the experimental environment in which it can operate. Recommendations to ISO from this community, the IAB and the vendors will carry great weight if they are in the language of the ISO common network management system and if they are rooted in actual experience with implementation and use in the field.)

- Responsibility for the SNMP effort should be placed in the hands of an IETF Working Group. (Rationale: Eliminate vendor-specific bias or control over the SNMP and its evolution and harmonize inputs from the Internet community.)

- As a high priority effort, define an extended Management Information Base (MIB) for SNMP and TCP/IP CMIP to bring them into closer conformance with the MIB defined for the experimental High Level Entity Management System (HEMS). (Rationale: The HEMS effort produced a very thorough and widely-discussed set of elements to monitor, along with definitions of the semantics of these elements. The current SNMP definitions are more restricted and the CMIP definitions less precise. Implementation of SNMP in a timely and useful fashion through the Internet cannot be satisfactorily completed without such a definition of information elements in hand.)

Recommendations

The ad hoc committee therefore recommends immediate action by the IAB on all four of these points. It should be noted that this resolution would not have been possible in such a timely way without the statesman-like efforts of Craig Partridge who, at the end of the day, recommended that the HEMS effort be withdrawn from consideration so as to pave the way for an Internet-wide agreement. In consideration of this unselfish act, the ad hoc committee urges the IAB to approve the recommendations above and to instruct the IETF to move quickly to accept and act on the SNMP items requiring completion.

[**Editor's Note**: A more detailed version of this article, including a comprehensive list of references, can be found in RFC 1052. In future issues of *ConneXions* we will be reporting on the progress of the 3 working groups; The MIB Working Group, The "Netman" Group, and the SNMP Extensions Working Group. As I write this, Advanced Computing Environments is hosting a Netman meeting where plans for a muti-vendor demonstration of interoperable Network Management is being discussed. This demonstration is scheduled to take place at the next TCP/IP Interoperability Conference/Exhibition, September 26 - 30, 1988.]

DoD Support of OSI Protocols

by Gladys Reichlen and Phill Gross, The Mitre Corp.

History

For years the DoD has been striving to implement a standard suite of computer networking protocols: namely those specified in the MIL-STD series and augmented by the applicable DARPA Request-For-Comments (RFC) documents. This, of course, is the well known TCP/IP Internet Protocol Suite that we all know and love. In 1985, in response to a report from the National Research Council, the DoD began evaluating the OSI international standard protocols. At that time, the OSI standards were not sufficiently developed to provide the support required for military needs. By late 1986, however, enough progress had been made for a Government OSI User's Group to be formed and, by April 1987, this group had developed the Government OSI Profile (GOSIP). GOSIP, which is based on the NBS OSI Implementor's Workshop Agreements, specifies the details for the use of the OSI protocols by the U.S. government. (Note: OSI standards are very flexible and generally have a large number of options. Implementor's Agreements are typically used to set such details as addressing format, option settings, and parameter values, in order to ensure interoperability between different implementations).

Following GOSIP, DoD issued a policy directive in July 1987 that clearly laid out a schedule for adoption of the OSI protocols as a military co-standard. Subsequently, DCA developed a strategy for introducing the usage of the OSI protocols into DoD networks. The resulting plan, entitled "The DoD OSI Implementation Strategy", has been reviewed by the DoD Protocol Standards Steering Group (PSSG) and is expected to be issued shortly. This article gives an overview of the main points of this plan.

Coexistence

The cornerstone of the DoD's OSI implementation strategy is coexistence and interoperability with the current TCP/IP Internet protocols. If the OSI protocols can operate simultaneously with the DoD protocols on the same backbone, and if provisions are made for these two suites to interoperate, then there does not need to be a disruptive "flag day" type of transition. Instead, transition from the DoD to OSI protocols can be an evolutionary process, guided at least in part by normal procurements.

The strategy

This strategy will be implemented by: (1) refining the set of OSI protocols and options specified by GOSIP for the DoD; (2) developing methods for DoD/OSI interoperability; and (3) encouraging the development of a suite of public domain software containing the complete OSI protocol stack.

GOSIP designates a pure OSI protocol stack (as opposed to any possible stacks consisting of a mixture of OSI and TCP/IP protocols at various layers) because that is what vendors are expected to most commonly provide. This means that interoperability between the protocol suites must be done by software at the Application Layer.

Timing for the DoD OSI implementation stratgy relies heavily on the July 1987 DoD directive. The table below shows the 6 target stages established for this migration.

Target Capability	*Features*
(1) Limited OSI Support	Basic X.25 Support, Early vendor products
(2) Experimental DoD/OSI Interoperability	Prototype Application Layer Gateways
(3) Limited DoD/OSI Interoperability	Operational Application Layer Gateways, Dual Protocol Hosts and Gateways
(4) OSI support equivalent to current TCP/IP protocols	OSI Routing, Virtual Terminal Protocol
(5) Advanced OSI Support	Network Management and Security Services
(6) Advanced DoD/OSI Interoperability	NATO Interoperability

Some specifics of the strategy include:

Application bridges

- *Selected Application Level Interoperability:*
DCA is developing full function application level bridges for file transfer and electronic mail. DDN subscribers will be supplied access to these application bridges as a service of the network. In related tasking, DCA is funding the development of prototype application layer bridges for both electronic mail and file transfer. These prototype application bridges will likely be used in at least the early stages of the transition.

IP level connectivity

- *Full Internetwork-Level OSI Connectivity between DoD Networks:*
DCA currently provides a set of internetwork gateways, called the Core Gateways. DCA will upgrade the Core Gateway system to support OSI protocols at the Internetwork Layer. This will eventually provide OSI connectivity equivalent to the current TCP/IP connectivity. Because development efforts supporting OSI routing protocols are still in the early stages, the use of DoD routing protocols is being investigated. However, initially manual methods may be required to perform routing table updates.

Iterim solutions

- *Other interim measures:*
While major computer vendors are committed to supporting the 7 layer OSI communications architecture, none can support it fully today. This is because key functionality at Layers 3 and 7 have not yet been fully approved as international standards. Efforts are underway to select interim methods to support those yet unstandardized functions such as Directory Services, Network Management and Security, and Internetwork Routing.

continued on next page

DoD Support of OSI Protocols *(continued)*

The full OSI suite

Software is being developed, as a cooperative effort by several universities, corporations and government agencies, which supports a complete OSI protocol profile in a POSIX-conformant environment. A major goal of this research effort, which is being coordinated by NBS, is the development of standard networking interfaces for computer operating systems. Complete details of this special project will be elaborated in a future issue of *ConneXions*. A by-product of this POSIX-based cooperative development effort will be the basic prerequisite software for application layer bridges and dual protocol gateways. This software package will be available to the DoD community for incorporation into their network planning.

Details remain to be worked out, particularly in the area of providing advanced requirements, such as security, full scale internet routing and network management. However, this approach allows the DoD community to begin using OSI protocols as they become available, while minimizing the interoperability limitations inherent in any transition period.

GLADYS A. REICHLEN is a Member of the Technical Staff at the Mitre Corporation, where she has been involved in program management and planning for data communications networks for the Defense Communications Agency. She is currently involved with both OSI transition and Internet management projects. Gladys received a BA in Mathematics from Trenton State University, Trenton, New Jersey.

PHILL GROSS is a Member of the Technical Staff at the Mitre Corporation, where he is involved in DoD protocol engineering, protocol performance and DoD OSI transition planning. He is chair of the Internet Engineering Task Force, whose charter it is to address short- to mid-term network and protocol architecture issues. His MS in Computer Science is from The Pennsylvania State University.

Conference and Exhibition Update

As mentioned in our April issue, the *3rd TCP/IP Interoperability Conference* will be held September 26-30, 1988 at the Santa Clara Convention Center and the Doubletree Hotel. Some 30 vendors have already signed up for the first *TCP/IP Interoperability Exhibition & Solutions Showcase* which runs concurrently with the conference. The exhibitors will all be connected to the "*Show and Tel-Net*" which is being managed for Advanced Computing Environments by The Wollongong Group.

In addition to showing TCP/IP interoperability, a group of vendors is also planning a demonstration of interoperable network management tools for TCP/IP networks based on the ISO CMIS/CMIP standard.

A conference program will be mailed to you in the near future. If you have any questions about exhibiting, please contact Margot Lockwood, Director of Conferences, at 415-941-3399.

Improving Your TCP: Look under the hood!

by Craig Partridge, BBN Laboratories Inc.

Recently my car has been acting up. Every few months it seems to need yet another trip to the mechanic to get yet another component fixed. Much of this repair is just normal wear and tear (Boston driving/roads can do horrible things to a car), but some of it could have been avoided if I'd been a little better about lifting the hood, seeing how things were working, and replenishing the fluids, or replacing the belts when I needed to. The purpose of this column is to suggest that your TCP implementation can also benefit from regular attention and maintenance. The implementation which ran just fine over smaller IP networks two years ago may be having severe performance problems on the larger, more congested networks of today.

Packet printer

The key maintenance tool for any TCP implementation is a *packet printer*. A packet printer is a program that prints out a copy of every packet that is sent and received over your TCP connection. Asking around will usually lead you to a packet printer somewhere on your network. If not, it is usually easy enough to jury-rig your TCP implementation to print out each packet it sends or receives along with a timestamp. Armed with your packet printer, try making TCP connections to various sites using anonymous FTP of medium or large files, and take a trace. Then sit down and read the trace carefully for odd behavior.

Protocol violations and bad behaviour

What defines odd behavior is a matter of taste. Any behavior that violates the TCP specification is clearly odd behavior, in fact, it is wrong behavior and should be repaired. One way to find protocol violations of this sort is to try some of the tests mentioned in Postel's RFC 1025, "TCP And IP Bake Off." But there is also a certain amount of behavior which, while within spec, is still poor. Numerous retransmissions are usually a sign of a bad implementation. Most networks have loss rates of only 1% or 2%. If your TCP retransmits more often than that, then in all likelihood, it has a poor round-trip time estimator or does not respond properly to congestion. Both problems have known fixes that should be applied. (see *ConneXions* Volume 1, No. 3 and No. 7)

Dying connections

Other more subtle problems may appear. One TCP implementation I know of would mysteriously die in the middle of a connection. This only happened occasionally, and there were no obvious network problems to explain it. A packet printer showed that, in some rare situations, the implementation got into a state where it never acknowledged new data packets. Another implementation sends reset packets in mid-connection for no obvious reason.

The point is, that a little attention to your TCP implementation may really pay off in better performance. As IP networks are getting larger, we are putting a strain on many TCP implementations. Attention now is likely to allow you to avoid problems later.

CRAIG PARTRIDGE received his B.A. from Harvard University in 1983, and has been a part-time Ph.D. candidate there since 1987. For the past five years he has worked for BBN Laboratories on a variety of networking related projects including CSNET, the NSF Network Service Center, and various projects concerned with distributed systems, IP transport protocols, and network management. He is a member of several Internet Task Forces.

CONNEXIONS
480 San Antonio Road
Suite 100
Mountain View, CA 94040

CONNEXIONS

PUBLISHER Daniel C. Lynch

EDITOR Ole J. Jacobsen